An Optimal Multi-System Control Measure Using the Approach of Conjugate Gradient Algorithm (CGA)

Journal:	*INFORMS Journal on Computing*
Manuscript ID:	JOC-2014-02-OA-027
Manuscript Type:	Original Article
Date Submitted by the Author:	04-Feb-2014
Complete List of Authors:	Adeleke, Olawale; Covenant University, Ota, Industrial Mathematics Adewumi, Aderemi; University of KwaZulu-Natal, Westville Campus., School of Mathematics, Statistics and Computer Science Omoregbe, Nicholas; Covenant auaniversity, Ota, Computer and Information Sciences Raji, Adekunle; Osun State Polytechnic, Ire, Mathematics
Keywords:	017 Computational, 019 Computational-experiments, 440 Mathematics, 452 Mathematics : Matrices, 682 Queues, Algorithms

An Optimal Multi-System Control Measure Using the Approach of Conjugate Gradient Algorithm (CGA)

Adeleke O. J[*1], , Adewumi O. A[2], Omoregbe N. I[1] and Raji A[3].

[*1] (Corresponding Author) Department of Computer and Information Science/Mathematics, Covenant University, Ota. Nigeria.
[2]School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Westville Campus, South Africa.
[3] Department of Mathematics and Statistics Osun State Polytechnic, Ire. Nigeria.

Abstract

In this paper we examine the application of the classical conjugate gradient method to queue theory. The parameters of the symmetric definite positive linear operator of a quadratic cost functional were obtained from the various characteristic features of a multi-channel queue system. The outcome was tested with numerical values and a comparison was made for systems with two, three and four service points. The numerical computations were carried out in a Maple 14 environment. The results obtained validate previous work done with a single-channel system.

Keywords: Conjugate gradient method, quadratic function, positive definite, matrix, symmetric, queue theory, multi-channel.

1. Introduction

The conjugate gradient method (CGM) was first introduced by Hestenes and Stiefel in 1952 [7]. The whole idea then was to develop a scheme for solving a system of linear equations. Later it was extended to nonlinear problems. It has since been a very powerful optimization tool for solving both the constrained and unconstrained optimization problems whether large or small scaled. Over the years, researchers have resorted to the use of CGM due to its flexibility and high rate of convergence.

In 2005, however, Omolehin et al [12] approached the concept of CGM from a new perspective by introducing the parameters of a CGM for quadratic functions from the classical queue theory. They considered the parameters of a single channel model to analyse the effectiveness of a CGM for a quadratic functional. It was evident that the queue system they considered has the following characteristics.

- Poisson arrival rate distribution
- Exponential service rate distribution
- An infinite population
- “first come, first serve” service rule

Such queue systems are based on some assumptions which make them work under virtually any given condition. It must also be assumed that items in line for service behave normally. In other words, they do not renege or balk under any circumstance. Also, there must exist a normal behaviour in the service machinery. This means the service facilities process each customers’ request continuously without a break and only one item is permitted to be served at a time.In this work, however, efforts have been directed to the multi-system where the number of service point is more than one.

As stated above, the classical linear CGM solves a quadratic function of the form

$$F(x) = F_0 + < b, x > + \frac{1}{2} < x, Lx > \qquad (1)$$

where L is a linear operator defined on a Hilbert space H. As we shall later show, L is symmetric and positive definite. The parameters of L were constructed from the characteristics of a multi-service point system. These parameters were carefully chosen so as to preserve the nature of L. If L is not positive definite, the descent in gradient of $F(x)$ is lost. It is also interesting to know that even if L is not symmetric positive definite (SPD), it can easily be amended to a SPD by multiplying the system by L^T, the transpose of L, to make it the required SPD.

There are several queuing models for which we would have constrained work, but we have been motivated to study the behaviour of a multi-service system because this will give us a clue to the behavioural pattern of several service points. A comparison was made for a two-, three-, and four-service points systems.

The remainder of this paper is organised as follows. In section 2, a brief review of queue theory and it applications were presented. In section 3, we investigated the various characteristics of a multi-service channel system. It was here that we arrived at the system parameters used in this work. We introduced the conjugate gradient method for a quadratic function in section 4. The conjugate gradient algorithm and the generated results were presented under section 5. Finally, in section 6, a brief remark was given on this work.

2. Queuing Theory

Queues are part of everyday life. Providing too much service requires overhead costs. Not doing so means a queue becomes excessively long. Thus in a nutshell, queue theory is a branch of study which addresses the problems associated with queues [11]. Queue models are generally stochastic models; although they could occasionally take the form of a deterministic model. Queue models ultimately aimed at studying the characteristics of a queue in order to achieve an economic balance between the cost of service and the cost associated with the waiting in line for service.

The theory on queue, as it were, has been applied to various decision making problems. In the health care sector for instance, the impact of bed assignment regulations was evaluated on time spent in queue, service utilisation and the effect of balking by McClain [10]. The problems of health workers roistering was studied by Cheang et al [2] while work on the combined effects of roistering and scheduling of staff was carried out by Ernst et al [3]. A brief compilation of works on the applications of queue theory in health care can be found in [13]. Rafaeli et al [14] of the Technion Institute of Technology (TIT), Israel, examined the impacts of queue structures on human behavioural attitude. The outcome of experiments carried out shows that customers nearest to a service facility are more pleased than those farther away. Queuing theory was applied to telecommunication networks [4] and a combine analysis with Markov process was applied to communication networks [1]. Newell in 1982 studied the flow of traffic at busy hours using queue theory. His proposed model was both efficient and effective [9].

However, the first work on queuing theory was published by A. K. Erlang in 1903 in a work on telephone traffic [5]. Generally, a queue has the following features: the length of the queue, system length, waiting time in queue, total time spent in a queue and the utilization factor. A system comprises of all items waiting to be served and those being served.

In a separate work by Kandall in 1903 and Lee in 1966 [5], the notations of a well-defined queue model were introduced. A typical notation of a queue model is

$$(u\ /\ v\ /\ w):(x\ /\ y\ /\ z)$$

where u is the item arrival distribution, v is item service distribution, w is the number of service point, x is the service discipline, y is the maximum number of items in a system and z is the source or population from which an item is drawn. Whether a system is single-channelled or multi-channelled depends on the value of w.

3. Quadratic Conjugate Gradient Method

In what follows, we have resorted to the use of usual notation for scalar product $< . \ , \ . >, \mathcal{H}$ is a Hilbert space and as noted earlier, L is a linear operator.

Given a quadratic functional of the form in (1) abovewhere $x, b \epsilon \mathcal{H}$, the basic problem which a conjugate gradient algorithm solves is to find a x^* which minimizes (1) such that

$$F(x^*) \leq F(x) \quad (2)$$

In doing this, two lines of actions are usually considered: (i) an initial guess of the minimizer x^* is made, say x_o. (ii) a sequence of points $[x_o, \ x_{1,} \ x_2, \ . \ . \ .]$ is constructed in a way that the condition

$$F(x_{m+1}) \leq F(x_m) \quad (3)$$

is satisfied. The sequence continues until no further x_{m+1} can be found to satisfy (3). At this point, the sequence approaches the minimizing argument of (1).

The conjugate gradient method is a conjugate descent method. By conjugate descent, it is assumed that there exists a sequence

$$\{d_m\}_{m=0}^{\infty} = [d_0, \ d_1, \ . \ . \ ., \ d_m, \ . \ . \ .] \quad (4)$$

where each of the numbers of $\{d_m\}$ is conjugate with respect to the linear operator L in (1). By this ii means

$$\langle d_n, \ Ld_m \rangle \{^{=0 \ if \ n \neq m}_{\neq 0 \ if \ n=m}$$

Since L is taken to be positive definite

$$\langle d_n, \ Ld_m \rangle > 0$$

Thus, by conjugate descent, a sequence

$$\{x_m\} = [x_0, \ x_1, \ . \ . \ ., \ x_m, \ . \ . \ .] \quad (5)$$

is constructed by guessing an initial point x_0 and setting

$$x_{m+1} = x_m + \alpha_m d_m \quad (6)$$

where α_m is a scalar constant also called the step length, and is chosen so that

$$F(x_m + \alpha_m d_m) = F_0 + < b, x_m + \alpha_m d_m > + \frac{1}{2} < x_m + \alpha_m d_m,\ L(x_m + \alpha_m d_m) >$$

(7) is minimized accordingly.

With the initial guess of minimizer made, the subsequent members of the sequence can be determined using the following relations [6]:

$$d_0 = -g_0 = -(b + Lx_0) \tag{8a}$$

$$x_{m+1} = x_m + \alpha_m d_m \tag{8b}$$

$$\alpha_m = \frac{<g_m, g_m>}{<d_m, Ld_m>} \tag{8c}$$

$$g_{m+1} = g_m + \alpha_m L d_m \tag{8d}$$

$$d_{m+1} = -g_{m+1} + \beta_m d_m \tag{8e}$$

$$\beta_m = \frac{<g_{m+1}, g_{m+1}>}{<g_m, g_m>} \tag{8f}$$

Where g_0 is the gradient of $F(x)$ at the initial point x_0 in the sequence (5),d_m is the initial search direction and β_m is the conjugate gradient parameter.

4. Multi-Channel Service System

Multi-channel queuing theory treats situations where there is several service points in parallel and each customer in the waiting line can be served by more than one service station.

For the purpose of this research, the arrival rate α and the sercice rate β are both mean values from the Poisson and exponential distributions respectively.

The various assumptions are:

$n =$ Number of customers in the system

$P_n =$ Probability of n customers in the system

$c =$ Number of service points ($c > 1$)

$\alpha =$ Arrival rate of customers

$\beta =$ Service rate of each channel

When $n > c$, there is no queue because all new arrivals are serviced instantly, and the rate of servicing will in this case be $n\beta$ as only n channels are working.

When $n = c$, all channels will be busy and when $n > c$, there will be $(n - c)$ customers in the queue and the rate of service for this instance will be $c\beta$.

In general, there are three cases of interest

I. When $n = 0$.

Considering a steady state

$P_1 = \frac{\alpha}{\beta} P_0$

II. When $1 \leq n \leq c - 1$.

In this case, for a steady state system

$$\alpha P_{n-1} - (\alpha + n\beta)P_n + (n+1)\beta P_{n+1} = 0 : 1 \leq n \leq c - 1$$

i.e.,$P_n = \frac{1}{n!}\left(\frac{\alpha}{\beta}\right)^2 P_0$

iii. When $n \geq c$

from (2),

$$\alpha P_{c-2} - [\alpha + (c-1)\beta]P_{c-1} + c\beta P_c = 0$$

this consequently results in

$$P_n = \frac{1}{c^{n-c}c!}\left(\frac{\alpha}{\beta}\right)^n P_0$$

From these the various properties of a multi-channel service system can be obtained.

The utility rate: $\rho = \frac{\alpha}{c\beta}$

Average number of items in the system: $S_i = \frac{\alpha\beta\left(\frac{\alpha}{\beta}\right)^c}{(c-1)!(c\beta-\alpha)^2} P_0 + \frac{\alpha}{\beta}$

Average number of item in queue: $Q_i = \frac{\alpha\beta\left(\frac{\alpha}{\beta}\right)^c}{(c-1)!(c\beta-\alpha)^2} P_0$

Average time spent in a system: $S_t = \frac{\beta\left(\frac{\alpha}{\beta}\right)^c}{(c-1)!(c\beta-\alpha)^2} P_0 + \frac{1}{\beta}$

Average queuing time: $Q_t = \frac{\beta\left(\frac{\alpha}{\beta}\right)^c}{(c-1)!(c\beta-\alpha)^2} P_0$

The probability that an item has to wait: $P_i(n \geq c) = \frac{\alpha\beta\left(\frac{\alpha}{\beta}\right)^c}{(c-1)!(c\beta-\alpha)} P_0$

The probability that an item enters the service without waiting:

$$P_{i-1}(n \geq c) = 1 - P_i(n \geq c)$$

The multi-channel service system described above which we also make use in this work is one that has an infinite item population. In the Kendall notation, this we represent as

$$(u\ /\ v\ /\ c) : (FCFS\ /\ \infty\ /\ \infty)$$

Where u and v are Poisson and Exponential distribution respectively, while $FCFS$ means First Come First Serve.

5. Main Results

Now we are set to present the main findings of this work. First, we recall from above that a CGM solves a quadratic function of the form $F(x) = F_0 + < b, x > + \frac{1}{2} < x, Lx >$, where L is a linear symmetric definite positive operator and $b, x \in \mathcal{H}$. F_0, the cost function has been fixed at zero. A fixed value of $b \epsilon \mathcal{H}$ have also been chosen. With the various multi-channel service system described above, the operator L have been constructed in the following way.

$$L = \begin{pmatrix} Q_i & U_r & S_i \\ U_r & Q_t & A_s \\ S_i & A_s & S_t \end{pmatrix}$$

The entries in the matrix above as discussed in section 4 are entered as follows

$$L = \begin{pmatrix} \frac{\alpha\beta\left(\frac{\alpha}{\beta}\right)^c}{(c-1)!\,(c\beta-\alpha)^2}P_0 & \frac{\alpha}{c\beta} & \frac{\alpha\beta\left(\frac{\alpha}{\beta}\right)^c}{(c-1)!\,(c\beta-\alpha)^2}P_0 + \frac{\alpha}{\beta} \\ \frac{\alpha}{c\beta} & \frac{\beta\left(\frac{\alpha}{\beta}\right)^c}{(c-1)!\,(c\beta-\alpha)^2}P_0 & c - \frac{\alpha\beta\left(\frac{\alpha}{\beta}\right)^c}{(c-1)!\,(c\beta-\alpha)^2}P_0 + 2\frac{\alpha}{\beta} \\ \frac{\alpha\beta\left(\frac{\alpha}{\beta}\right)^c}{(c-1)!\,(c\beta-\alpha)^2}P_0 + \frac{\alpha}{\beta} & c - \frac{\alpha\beta\left(\frac{\alpha}{\beta}\right)^c}{(c-1)!\,(c\beta-\alpha)^2}P_0 + 2\frac{\alpha}{\beta} & \frac{\beta\left(\frac{\alpha}{\beta}\right)^c}{(c-1)!\,(c\beta-\alpha)^2}P_0 + \frac{1}{\beta} \end{pmatrix}$$

This reduces to

$$L = \begin{pmatrix} \frac{\alpha\beta}{\tau}P_0 & \frac{\alpha}{c\beta} & \frac{\alpha\beta}{\tau}P_0 + \frac{\alpha}{\beta} \\ \frac{\alpha}{c\beta} & \frac{\beta}{\tau}P_0 & c - \frac{\alpha\beta}{\tau}P_0 + 2\frac{\alpha}{\beta} \\ \frac{\alpha\beta}{\tau}P_0 + \frac{\alpha}{\beta} & c - \frac{\alpha\beta}{\tau}P_0 + 2\frac{\alpha}{\beta} & \frac{\beta}{\tau}P_0 + \frac{1}{\beta} \end{pmatrix}$$

where

$$\tau = \frac{\left(\frac{\alpha}{\beta}\right)^c}{(c-1)!\,(c\beta-\alpha)^2}$$

This system, being not real comes with incertitude. Suppose this incertitude parameter is introduced as δ only in the first entry of the matrix, this further modifies the matrix to

$$L = \begin{pmatrix} \delta P_0 & \frac{\alpha}{c\beta} & \frac{\alpha\beta}{\tau}P_0 + \frac{\alpha}{\beta} \\ \frac{\alpha}{c\beta} & \frac{\beta}{\tau}P_0 & c - \frac{\alpha\beta}{\tau}P_0 + 2\frac{\alpha}{\beta} \\ \frac{\alpha\beta}{\tau}P_0 + \frac{\alpha}{\beta} & c - \frac{\alpha\beta}{\tau}P_0 + 2\frac{\alpha}{\beta} & \frac{\beta}{\tau}P_0 + \frac{1}{\beta} \end{pmatrix}$$

With this in place, the quadratic function in (1) was solved by implementing the following algorithm using a Maple code.

Step 1. Set $m = o$ and choose x_o and compute $g_0 = b + Lx_0 = \nabla f(x_0)$. If $g_0 = 0$, stop. Else set $d_0 = -g_0$.

Step 2. Set $\alpha_m = \frac{<g_m,\ g_m>}{<d_m,\ Ld_m>}$, compute $x_{m+1} = x_m + \alpha_m d_m$ and $g_{m+1} = g_m + \alpha_m L d_m$. If $g_{m+1} = 0$, stop.

Step 3. Set $\beta_m = \frac{<g_{m+1}, g_{m+1}>}{<g_m, g_m>}$ and compute $d_{m+1} = -g_{m+1} + \beta_m d_m$.

Step 4: set $m = m + 1$ and go to step 2.c

The results obtained are presented in the following tables. The variables used are $\alpha = 6$, $\beta = 4$ and $P_0 = 0.2$.

Table 1a. $c = 2$

Incertitude Parameter	x^*	$\|g\|$	f^*
$\delta = 1$	$x_1 = -2.434513$	1.565334e-08	-0.562630
	$x_2 = 6.962669$	1.087638e-08	
	$x_3 = 0.301167$	1.005755e-08	

Table 1b. $c = 2$

Incertitude Parameter	x^*	$\|g\|$	f^*
$\delta = 0.5$	$x_1 = -2.220554$	9.417628e-10	-0.832928
	$x_2 = 6.569801$	1.812216e-09	
	$x_3 = 0.308263$	4.228760e-09	

Table 1c. $c = 2$

Incertitude Parameter	x^*	$\|g\|$	f^*
$\delta = 2$	$x_1 = -3.015654$	5.912982e-09	0.171535
	$x_2 = 8.029749$	4.579172e-09	
	$x_3 = 0.281893$	7.868555e-09	

Table 2a. $c = 3$

Incertitude Parameters	x^*	$\|g\|$	f^*
$\delta = 1$	$x_1 = -10.426762$	5.464351e-09	15.823238
	$x_2 = 15.060572$	7.547903e-09	
	$x_3 = 0.132159$	1.271628e-08	

Table 2b. $c = 3$

Incertitude Parameters	x^*	$\|g\|$	f^*
$\delta = 0.5$	$x_1 = -8.075488$	1.639970e-08	11.613178
	$x_2 = 12.056722$	1.125223e-08	
	$x_3 = 0.185521$	1.090176e-08	

Table 2c. $c = 3$

Incertitude Parameters	x^*	$\|g\|$	f^*
$\delta = 2$	$x_1 = -24.963744$	2.323269e-08	41.852340
	$x_2 = 33.632169$	2.197065e-08	
	$x_3 = -0.197759$	2.674142e-08	

Table 3a. $c = 4$

Incertitude Parameters	x^*	$\|g\|$	f^*
$\delta = 1$	$x_1 = -40.094739$	1.396615e-08	80.542416
	$x_2 = 40.604463$	1.357060e-08	
	$x_3 = -0.287554$	1.159818e-08	

Table 3b. $c = 4$

Incertitude Parameters	x^*	$\|g\|$	f^*
$\delta = 0.5$	$x_1 = -20.351863$	2.718575e-09	39.742283
	$x_2 = 21.271653$	4.537117e-09	
	$x_3 = 0.013885$	2.819973e-08	

Table 3c. $c = 4$

Incertitude Parameters	x^*	$\|g\|$	f^*
$\delta = 2$	$x_1 = 42.646986$	1.278466e-08	-90.449562
	$x_2 = -40.418692$	1.026607e-08	
	$x_3 = 0.975765$	6.391723e-08	

6. Remarks

In this computation, the tolerance was taken as $1.0e - 06$. In table 1, for a two-channel service system, it was observed that there was a successive decrease in the gradient norm when the uncertainty parameter, $\delta = 1$. However, this was not the case when $\delta = 0.5$ and $\delta = 2$. For the three-channel service system, that is when $c = 3$, this successive decrease only occurs at $\delta = 0.5$. For$c = 4$, it occurs at $\delta = 1$. Our observation is however that a good approximation as we have in table 1a, 2b and 3a will be achievable through a careful selection of δ. A good choice will almost likely be $0.5 < \delta < 1$. This validates the choice of $p = 0.999 < 1$ in the work of Omolehin et al [12].

7. References

1. Busson, A. 2002. Markov chain, markov processes, and queuing theory and application to communication networks.
2. Cheang, B., Lib, H., Lim, A., and Rodrigues, B. 2003. Nurse roistering problems: a bibliographic survey. European Journal of Operation Research.
3. Ernst, A. T., Jiang, H., Krishnamoorthy, M., and Sier, D. 2004. Staff scheduling and roistering: a review of applications, methods and models. European Journal of Operation Research. 153 (1), pp. 3-27.
4. Giambene, G. Queuing theory and its applications in the telecommunications networks. Kluwer Academic Publishers.
5. Gupta, P. K. and Hira, D. S. 2011. Operations research. S. Chand. 6th edition, pp. 916.
6. Hasdorff, L. 1976. Gradient optimization and nonlinear control. John Wiley, New York.
7. Hestenes, M.R. and Stiefel, E. 1952. Method of conjugate gradient for solving linear equations. J. Res. Nat. Bur. Stand. 49.
8. Lakshmi, C. and Iyer, S. A. 2013. Application of queuing theory in health care: a literature review. Operations Research for Health Care, vol. 2.
9. Marin, C. V., Drury, C. G., Batta, R. and Lin, L. 2007. Human factors contribute to Queuing theory: Parkinson's Law and security screening. Proceedings of the Human Factors and Ergonomics Society, 51st Annual Meeting.
10. McClain, J. O. 1976. Bed planning using queuing theory models of hospital occupancy: a sensitivity analysis. Inquiry, 19, pp. 167-176.
11. Morse, P. M. 1958. Queues, inventories and maintenance. Wiley, New York.

12. Omolehin, J. O., Rauf, K. and Evans, D. J. 2005. Conjugate gradient method approach to queue theory. International Journal of Computer Mathematics, 82:7, pp. 829-835.
13. Preater, J. 2002. Queues in health. Health care management sciences, 5, pp. 283.
14. Rafaeli, A., Barron, G. and Haber, K. 2002. The effects of queue structure on attitudes. Journal of Service Research, vol. 5, pp. 125.